

Tematy projektów z Metod Nieparametrycznych

Zad. 1

Za pomocą odpowiednich symulacji porównać moce testu Shapiro-Wilka, testu Kołmogorowa i testu χ^2 (z hipotezą główną, że cecha ma rozkład normalny) w przypadku, gdy dane pochodzą z rozkładu t -Studenta lub innej klasy rozkładów symetrycznych (uwaga: żeby wyniki były porównywalne trzeba pamiętać o standaryzacji danych). Uzyskane wyniki należy przedstawić na odpowiednich wykresach ilustrujących moce testów z uwzględnieniem:

- liczby danych,
- liczby stopni swobody generowanego rozkładu t -Studenta lub innych parametrów wybranego rozkładu.

Zad. 2

Za pomocą odpowiednich symulacji zbadać odsetek odrzuceń prawdziwej hipotezy głównej (tzn. błąd I rodzaju) w teście Kołmogorowa i teście χ^2 w przypadku weryfikacji zgodności z rozkładem χ^2 o różnych stopniach swobody, tzn. generować dane pochodzące z rozkładu χ^2 o k stopniach swobody i weryfikować hipotezę, że dane mają rozkład χ^2 o k st. swobody. Należy porównać wyniki klasycznego sposobu testowania i testowania z wykorzystaniem PIT (probability integral transform). Uzyskane wyniki należy przedstawić na odpowiednich wykresach ilustrujących moce testów z uwzględnieniem:

- liczby danych,
- liczby stopni swobody generowanego rozkładu.

Zad. 3

Za pomocą odpowiednich symulacji zbadać moc testu Kołmogorowa-Smirnowa w przypadku, gdy rozważane próbki danych pochodzą z rozkładów χ^2 o różnej liczbie stopni swobody. Do celów symulacji proszę przesunąć generowane dane tak, by miały tę samą wartość oczekiwaną. Uzyskane wyniki należy przedstawić na odpowiednich wykresach ilustrujących moce testów z uwzględnieniem:

- liczby danych (można rozważać różne liczebności w każdej z próbek),
- liczby stopni swobody.

Zad. 4

Za pomocą odpowiednich symulacji zbadać wpływ liczby klas (na które dzielone są dane) odsetek odrzuceń prawdziwej hipotezy głównej (tzn. błąd I rodzaju) w teście χ^2 w przypadku weryfikacji zgodności z rozkładem χ^2 o różnych stopniach swobody, tzn. generować dane pochodzące z rozkładu χ^2 o k stopniach swobody i weryfikować hipotezę, że dane mają rozkład χ^2 o k st. swobody. Uzyskane wyniki należy przedstawić na odpowiednich wykresach ilustrujących rozmiary testów z uwzględnieniem:

- liczby danych,

- liczby klas,
- liczby stopni swobody generowanego rozkładu.

Zad. 5

Za pomocą odpowiednich symulacji zbadać odsetek odrzuceń prawdziwej hipotezy głównej (tzn. błąd I rodzaju) w teście Kołmogorowa i teście χ^2 w przypadku weryfikacji zgodności z rozkładem t -Studenta o różnych stopniach swobody, tzn. generować dane pochodzące z rozkładu t -Studenta o k stopniach swobody i weryfikować hipotezę, że dane mają rozkład t -Studenta o k st. swobody. Należy porównać wyniki klasycznego sposobu testowania i testowania z wykorzystaniem PIT (probability integral transform). Uzyskane wyniki należy przedstawić na odpowiednich wykresach ilustrujących rozmiary testów z uwzględnieniem:

- liczby danych,
- liczby stopni swobody generowanego rozkładu.

Zad. 6

Za pomocą odpowiednich symulacji zbadać moc testu Kołmogorowa, testu Lillieforsa i testu Andersona-Darlinga (z hipotezą główną, że cecha ma rozkład normalny) w przypadku, gdy dane pochodzą z rozkładu t -Studenta (uwaga: żeby wyniki były porównywalne trzeba pamiętać o standaryzacji danych). Uzyskane wyniki należy przedstawić na odpowiednich wykresach ilustrujących moce testów z uwzględnieniem:

- liczby danych,
- liczby stopni swobody generowanego rozkładu t -Studenta.

Zad. 7

Za pomocą odpowiednich symulacji zbadać moc testu Kołmogorowa-Smirnowa w przypadku, gdy rozważane próbki danych pochodzą z rozkładów normalnych o tej samej wartości oczekiwanej, ale o różnych odchyleniach standardowych. Rozważyć różne wartości oczekiwane. Uzyskane wyniki należy przedstawić na odpowiednich wykresach ilustrujących moce testów z uwzględnieniem:

- liczby danych (można rozważać różne liczebności w każdej z próbek),
- odchyłeń standardowych.

Zad. 8

Za pomocą odpowiednich symulacji zbadać moc testu Kołmogorowa i testu Kołmogorowa z wykorzystaniem PIT (probability integral transform), z hipotezą główną, że cecha ma rozkład normalny w przypadku, gdy dane pochodzą z rozkładu t -Studenta lub z rozkładu χ^2 (uwaga: żeby wyniki były porównywalne trzeba pamiętać o standaryzacji danych). Uzyskane wyniki należy przedstawić na odpowiednich wykresach ilustrujących moce testów z uwzględnieniem:

- liczby danych,
- liczby stopni swobody generowanego rozkładu t -Studenta, χ^2 .

Zad. 9

Za pomocą odpowiednich symulacji zbadać odsetek odrzuceń prawdziwej hipotezy głównej (tzn. błąd I rodzaju) w teście Kołmogorowa w przypadku weryfikacji zgodności z rozkładem NIG oraz z rozkładem gamma o różnych wartościach parametrów. Należy porównać wyniki klasycznego sposobu testowania i testowania z wykorzystaniem PIT (probability integral transform). Uzyskane wyniki należy przedstawić na odpowiednich wykresach ilustrujących moce testów z uwzględnieniem:

- liczby danych,
- parametrów generowanego rozkładu.

Zad. 10

Za pomocą odpowiednich symulacji zbadać wpływ liczby klas (na które dzielone są dane) odsetek odrzuceń prawdziwej hipotezy głównej (tzn. błąd I rodzaju) w teście χ^2 w przypadku weryfikacji zgodności z rozkładem t -Studenta o różnych stopniach swobody, tzn. generować dane pochodzące z rozkładu t -Studenta o k stopniach swobody i weryfikować hipotezę, że dane mają rozkład t -Studenta o k st. swobody. Uzyskane wyniki należy przedstawić na odpowiednich wykresach ilustrujących rozmiary testów z uwzględnieniem:

- liczby danych,
- liczby klas,
- liczby stopni swobody generowanego rozkładu.

Zad. 11

Za pomocą odpowiednich symulacji zbadać moc testów: Jarque-Bera, Shapiro-Wilka i Lillieforsa w przypadku, gdy dane pochodzą z rozkładu t -Studenta. Uzyskane wyniki należy przedstawić na odpowiednich wykresach ilustrujących moce testów z uwzględnieniem:

- liczby danych,
- liczby stopni swobody generowanego rozkładu t -Studenta.

Zad. 12

Za pomocą odpowiednich symulacji zbadać moc testu Kołmogorowa-Smirnowa w przypadku, gdy rozważane próbki danych pochodzą z rozkładów normalnych o tym samym odchyleniu standardowym, ale różniących się wartościami oczekiwanymi. Rozważyć różne wartości odchyleń standardowych. Uzyskane wyniki należy przedstawić na odpowiednich wykresach ilustrujących moce testów z uwzględnieniem:

- liczby danych (można rozważać różne liczebności w każdej z próbek),
- wartości oczekiwanych.

Zad. 13

Za pomocą odpowiednich symulacji zbadać odsetek odrzuceń prawdziwej hipotezy głównej w teście Kołmogorowa-Smirnowa w przypadku, gdy rozważane próbki pochodzą z tego samego rozkładu. Rozważyć dwa różne rozkłady asymetryczne. W badaniu uwzględnić różną liczebność próbek.

Uzyskane wyniki należy przedstawić na odpowiednich wykresach ilustrujących wyniki z uwzględnieniem:

- liczby danych (można rozważać różne liczebności w każdej z próbek),
- parametrów rozważanych rozkładów.

Zasady oceniania

Projekt należy wykonać w programie R. Wykonanie projektu obejmuje: opis (w języku polskim) przeprowadzonych badań wraz z wynikającymi z nich wnioskami oraz skrypt w R użyty do przeprowadzenia potrzebnych obliczeń i utworzenia wykresów. Na ocenę pracy mają wpływ: zawartość merytoryczna pracy (75%), strona edytorska i logiczność opisu (25%).